Chris Piech
CS109

Fall 2021
Oct 26th

CS109 Midterm Exam

This is a closed calculator/computer exam. You are, however, allowed to use notes in the exam. You have 2 hours (120 minutes) to take the exam. The exam is 120 points, meant to roughly correspond to one point per minute of the exam. You may want to use the point allocation for each problem as an indicator for pacing yourself on the exam.

In the event of an incorrect answer, any explanation you provide of how you obtained your answer can potentially allow us to give you partial credit for a problem. For example, describe the distributions and parameter values you used, where appropriate. It is fine for your answers to include summations, products, factorials, exponentials, and combinations.

You can leave your answer in terms of Φ (the CDF of the standard normal). For example $\Phi(\frac{3}{4})$ is an acceptable final answer..

I acknowledge and accept the letter and spirit of the honor code. I pledge to write more neatly than I have in my entire life:

Signature: ______________________________

Family Name (print): ______________________________

Given Name (print): ______________________________

Email (preferably your gradescope email): ____________________

1 Color.com [20 points]

For covid testing, color.com produces 6 digit tracking codes. Code can start with 0s. Here are a few example codes: 123456, 000000, 091923.

a. (2 points) How many unique codes are possible with 6 digits?

10^6

There are 10 possible values that each digit can take on, $[0,9]$. We know each tracking code must always be *exactly* 6 digits long. Importantly, repetitive digit assignments are allowed, so for every placeholder we have 10 possible options. In other words, the number of possible values is $10*10*10*10*10*10 = 10^6$.

b. (5 points) Color.com generates 3 random codes using a bad algorithm: each code is generated uniformly over all possible codes without checking for duplicates. What is the probability that none of them 3 random codes are the same?

Solution 1: There are $(10^6)^3$ ways to generate 3 random codes since there are 10^6 possible options for each code — this is the sample space. The event space only requires 10^6 options for the first code, 10^6-1 options for the second code (the remaining number of non-duplicate options for the second code), and 10^6-2 options for the third code (the remaining number of non-duplicate options for the third code). Hence, the probability that none of the 3 randomly generated codes is a duplicate is

$$\frac{(10^6)\cdot(10^6-1)\cdot(10^6-2)}{(10^6)^3}$$

Solution 2: There are $(10^6)^3$ ways to generate 3 random codes since there are 10^6 possible options for each code — this is the sample space. There are $\binom{10^6}{3}$ ways to select 3 unique, non-duplicate codes from the set of passwords and there are 3! permutations for any set of 3 unique codes so the sample space is $3!\cdot\binom{10^6}{3}$. Hence, the probability that none of the 3 randomly generated codes is a duplicate is

$$\frac{3!\cdot\binom{10^6}{3}}{(10^6)^3}$$

Solution 3: We can evaluate the probability of the complement of the event (E) where none of the 3 random codes are duplicates:

$$\begin{aligned}P(E) &= 1-P(E^C)\\ &= 1-P(\text{two of the passwords are duplicates})-P(\text{all three are duplicates})\\ &= 1-\binom{3}{2}\cdot\left(\frac{10^6}{10^6}\right)\cdot\left(\frac{1}{10^6}\right)\cdot\left(\frac{10^6-1}{10^6}\right)-\left(\frac{10^6}{10^6}\right)\cdot\left(\frac{1}{10^6}\right)^2\end{aligned}$$

c. (5 points) Every time a user enters a digit, there is a 1% chance that they enter it incorrectly. A code is incorrect if any digit is incorrect. What is the probability of an incorrectly entered 6-digit-code?

Each digit has a $p = 0.01$ chance of being entered incorrectly, independently of other digits. Since there are a total of 6 digits to enter, the total number of incorrectly entered digits X is a binomial random variable, $X \sim Bin(n = 6, p = 0.01)$. Entering an incorrect code corresponds to entering at least 1 incorrect digit, the probability of which is

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{6}{0} 0.01^0 \cdot (1 - 0.01)^6 = 1 - 0.99^6$$

d. (8 points) 20,000 Stanford students enter their 6-digit-code independently. Use an approximation that could be used to efficiently compute the probability that more than 10 students will incorrectly enter a code.

The intuition to approach this question is to use a Binomial random variable with $n = 20000$ and $p = 1 - 0.99^6 \approx 0.0585$, which is the answer to 1c. Then we could approximate this Binomial random variable in two ways.

1. Since n is large and p is small, we can estimate it with $X \sim \text{Poisson}(\lambda = np)$. Therefore, $P(X > 10) = 1 - P(X \leq 10) = 1 - \Sigma_{i=0}^{10} \frac{(np)^i * e^{-np}}{i!}$
2. We could also estimate it using $Y \sim \text{Normal}(np, np(1-p))$. In this case, $P(Y > 10) = 1 - P(Y \leq 10) = 1 - \phi(\frac{10.5 - np}{\sqrt{np(1-p)}})$. Note that 10.5 on the numerator comes from the continuity correction.

2 Doodle Poll [20 points]

You are trying to find a time when you can meet in a group. Assume a day is composed of 8 one hour blocks. Each block has $p = \frac{5}{8}$ probability of being busy for each person. We assume that whether each block is busy for a person is independent of whether other blocks are busy for them, and also independent of whether that block is busy for other people.

a. (5 points) What is the probability that a single person is busy for 6 or more blocks in a single day?

Let B be the number of blocks that are busy. $B \sim$ Bin(8, 5/8)

$P(B >= 6) = \binom{8}{6} * (\frac{5}{8})^6 * (\frac{3}{8})^2 + \binom{8}{7} * (\frac{5}{8})^7 * (\frac{3}{8})^1 + \binom{8}{8} * (\frac{5}{8})^8 * (\frac{3}{8})^0$

Common mistakes:

- Using a Poisson or Normal approximation. N is not sufficiently large, p is not sufficiently small, and variance is not sufficiently large for these to be good approximations. See lecture 10 slide 62.
- Forgetting the $\binom{8}{k}$ term. You need to account for ways of permuting 6, 7, and 8 busy blocks from the 8 available blocks.

b. (7 points) Two people are trying to schedule a meeting. What is the probability that there is a block of time when they are both free on a given day?

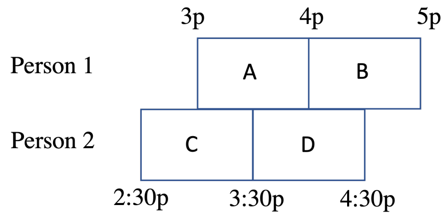
The probability that both the people are free in a given block is $(1 - p_{\text{one free}})^2$. Thus the probability that both are free is $p_{\text{both free}} = (\frac{3}{8})^2$

Let X be the number of blocks where they are both free. $X \sim Bin(n, p_{\text{both free}})$. We want the probability that there is a block when both are free so we want to find: 1 - P(X = 0). Which is

$$1 - \binom{8}{0}(1 - p_{\text{both free}})^8$$

$$1 - (1 - (\frac{3}{8})^2)^8$$

c. (8 points) A person in India and a person in the UK are trying to schedule a one-hour meeting. In India, time zones are offset by half an hour relative to time zones in the UK. Here are their overlapping blocks:



The two people can only meet if blocks A, C, and D are all free, or if blocks A, B, and D are all free. What is the probability that the two people will be able to meet? [1]

The 1st way: There are 3 possible cases for the two people to be able to meet: 1) A, B, D are free, and C is not free; 2) A, C, D are free, and B is not free; 3) A, B, C, D are all free. Since these three cases are mutually exclusive, we simply add their probabilities together to get the final answer:

$$(\frac{5}{8})(\frac{3}{8})^3 + (\frac{5}{8})(\frac{3}{8})^3 + (\frac{3}{8})^4 = \frac{351}{4096} \approx 0.0857$$

The 2nd way: Let's define two events. Event X: A, C, D are all free; and Event Y: A, B, D are all free. If we add their probabilities together, we will overcount the probability of A, B, C, D are all free which is actually $P(X \cap Y)$. Therefore, we can also calculate the final answer in this way:

$$(\frac{3}{8})^3 + (\frac{3}{8})^3 - (\frac{3}{8})^4 = \frac{351}{4096} \approx 0.0857$$

The 3rd way: By observation, A and D must be free for the two people to be able to meet, and after that, it is sufficient to have either B or C is free, which is equivalent to $1 - P(\text{B and C are not free})$. Therefore, we can also calculate the final answer in this way:

$$(\frac{3}{8})^2 * (1 - (\frac{5}{8})^2) = \frac{351}{4096} \approx 0.0857$$

[1] In case you are curious: all times in the figure are in India Standard Time. Person 2 is in the UK and can't meet before 2:30p because it would be too early in their local time.

3 I Heard That! [30 points]

We are trying to build a probabilistic at-home test to detect whether an 8 month old can hear a particularly soft sound. To do so, we are going to play the sound and then observe how much the baby's gaze moves in the following 3 seconds. If the baby hears the sound, they are more likely to have a large change in gaze. We have built a special camera which can record the change in gaze. Before observing any change in gaze, our prior belief is that:

$$\text{P(can hear the sound)} = \frac{3}{4}$$

X is our measure of the angle that the baby's gaze moves over 3 seconds. We have estimated the following probability mass functions for X. These PMFs were obtained by observing tens of thousands of babies with known hearing ability. All values are discretized into buckets covering 5 degrees:

Value of X	PMF of X given Baby can hear the sound	PMF of X given Baby can **not** hear the sound
0 to 5	0.08	0.40
5 to 10	0.15	0.30
10 to 15	0.35	0.12
15 to 20	0.20	0.08
20 to 25	0.12	0.05
Above 25	0.10	0.05

a. (8 points) You play a sound and observe no movement ($X = 0$). What is the updated probability that she can hear the sound?

Let Y be the event that the baby hears the sound. We are trying to calculate $P(Y|X=0)$. Using the definition of conditional probability and Bayes' Theorem, we know:

$$P(Y|X=0) = \frac{P(Y, X=0)}{P(X=0)} = \frac{P(X=0|Y)P(Y)}{P(X=0)}$$

We can then use the law of total probabilities to expand the denominator.

$$P(Y|X=0) = \frac{P(X=0|Y)P(Y)}{P(X=0|Y)P(Y) + P(X=0|Y^C)P(Y^C)}$$

From the provided values we know:

$$P(X=0|Y) = 0.08$$
$$P(Y) = 0.75$$
$$P(X=0|Y^C) = 0.40$$
$$P(Y^C) = 1 - P(Y) = 1 - 0.75 = 0.25$$

Plugging all these values in gives,

$$P(Y|X=0) = \frac{0.08 * 0.75}{0.08 * 0.75 + 0.40 * 0.25} = \frac{3}{8}$$

b. (6 points) You repeat the experiment to get a **second** independent reading and observe a 16 degree change — so you now have two independent observations, 0 degrees and 16 degrees. What is your newly updated probability? Let p be your answer to part (a).

We will once again update the probability of Y given this new reading. Using the definition of conditional probability and Bayes' Theorem, we know:

$$P(Y|X=16) = \frac{P(Y,X=16)}{P(X=16)} = \frac{P(X=16|Y)P(Y)}{P(X=16)}$$

We can again use the law of total probabilities to expand the denominator.

$$P(Y|X=16) = \frac{P(X=16|Y)P(Y)}{P(X=16|Y)P(Y)+P(X=16|Y^C)P(Y^C)}$$

From the provided values we know:

$$P(X=16|Y) = 0.2$$
$$P(Y) = p$$
$$P(X=16|Y^C) = 0.08$$
$$P(Y^C) = 1-P(Y) = 1-p$$

Note that $P(Y) = p$ because in part a we calculated the updated probability of Y given our $X = 0$ reading and $P(Y)$ is now equal to $P(Y|X = 0)$.

Plugging all these values in gives,

$$P(Y|X=16) = \frac{0.02*p}{0.02*p+0.08*(1-p)} = \frac{3}{5}$$

The Normal Assumption: We choose to approximate eye movements with normal distributions. For babies who **can** hear sounds, we approximate their gaze movement after the sound is played as: $N(\mu = 15, \sigma^2 = 50)$. For babies who can **not** hear sounds, we approximate gaze movement as $N(\mu = 8, \sigma^2 = 50)$.

c. (8 points) For a new baby we observe a 0 degree movement after the sound is played. What is your belief that a baby can hear, under The Normal Assumption? Since this is a new baby, you should not carry forward results from part (a) or (b).

Let X be the degree movement of a baby after the sound is played.
Let Y be the event that a baby can hear (i.e. Y = 1 if a baby can hear, Y = 0 if a baby cannot).
So from the question, we get: $P(X|Y=1) \sim N(\mu = 15, \sigma^2 = 50)$; $P(X|Y=0) \sim N(\mu = 8, \sigma^2 = 50)$.
We want to solve for: $P(Y=1|X=0)$.
First, apply the continuous variant of the Bayes' Theorem:

$$P(Y=1|X=0) = \frac{f_{X|Y=1}(0) \cdot P(Y=1)}{P(X=0)}$$

Since $P(X=0)$ is unknown, apply the law of total probabilities:

$$P(Y=1|X=0) = \frac{f_{X|Y=1}(0) \cdot P(Y=1)}{f_{X|Y=1}(0) \cdot P(Y=1) + f_{X|Y=0}(0) \cdot P(Y=0)}$$

Then, find all f's in the above formula. There can be two possible solutions from now on: 1. Plug in the respective μ and σ to the Normal PDF: $\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$. 2. Convert the non-standard Normal distributions to standard normal distributions (See below).
Method 1:

$$f_{X|Y=1}(x) = \frac{1}{\sqrt{50}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-15}{\sqrt{50}})^2}$$

Method 2: Let $Z = \frac{X-15}{\sqrt{50}}$, so $X = 15 + \sqrt{50}Z$. We know $Z_1 \sim N(0,1)$.

$$f_{X|Y=1}(x) = f(X=x|Y=1) = f(15+\sqrt{50}Z = x|Y=1) = f(Z = \frac{x-15}{\sqrt{50}}|Y=1) = \phi(\frac{x-15}{\sqrt{50}})$$

Find all components of $P(Y=1|X=0)$ using either method to get the following answers (both are accepted):
Method 1:

$$P(Y=1|X=0) = \frac{\frac{1}{\sqrt{50}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{0-15}{\sqrt{50}})^2} \cdot \frac{3}{4}}{\frac{1}{\sqrt{50}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{0-15}{\sqrt{50}})^2} \cdot \frac{3}{4} + \frac{1}{\sqrt{50}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{0-8}{\sqrt{50}})^2} \cdot \frac{1}{4}}$$

Method 2:

$$P(Y=1|X=0) = \frac{\phi(\frac{0-15}{\sqrt{50}}) \cdot \frac{3}{4}}{\phi(\frac{0-15}{\sqrt{50}}) \cdot \frac{3}{4} + \phi(\frac{0-8}{\sqrt{50}}) \cdot \frac{1}{4}}$$

d. (8 points) We want to check if The Normal Assumption was accurate. Write an expression for the probability that $10 < X < 15$ given that a baby **can hear the sound**, under the Normal Assumption. You can leave your answer in terms of Φ.

Similar to method 2 in 3c. Let $Z = \frac{X-15}{\sqrt{50}}$. Then $X = 15 + \sqrt{50}Z$.
When $X \in (10, 15)$, $Z \in (\frac{10-15}{\sqrt{50}}, \frac{15-15}{\sqrt{50}})$.

So,

$$P(10 < X < 15) = P(\frac{10-15}{\sqrt{50}} < Z < \frac{15-15}{\sqrt{50}}) = \Phi(\frac{15-15}{\sqrt{50}}) - \Phi(\frac{10-15}{\sqrt{50}})$$

4 Sleep and Dreams [30 points]

The number of times a baby wakes up in the night (an 8 hour period) is Poisson with rate lambda = 2.

a. (6 points) What is the probability that the baby wakes up exactly once during the night?

Let X be the number of times a baby wakes up during the night. By the question, one has $X \sim \text{Poi}(2)$. Therefore, by PDF for Poission distribution, one has

$$P(X=1) = \frac{\lambda^1 e^{-\lambda}}{1!} = \frac{2e^{-2}}{1!} = 2e^{-2}$$

b. (6 points) The amount of rest you get is 0 if the baby wakes up more than 3 times. Otherwise, the amount of rest you get is $15-2X$, where X is the number of times the baby wakes up. What is the expectation of rest?

Let R be the amount of rest one gets, and let X be the number of times the baby wakes up as described in the question. One can see that in this case, one has

$$R(x) = \begin{cases} 15-2x & x \leq 3 \\ 0 & x > 3 \end{cases}$$

Therefore, the expectation of R can be computed as below:

$$\begin{aligned} \mathbb{E}[R] &= \sum_{\text{possible values r for R}} r \cdot P(R=r) \\ &= \sum_{\text{possible values x for X}} R(x) \cdot P(X=x) \\ &= \left(\sum_{x\leq 3} (15-2x) \cdot P(X=x)\right) + \left(\sum_{x>3} 0 \cdot P(X=x)\right) \\ &= \sum_{x\leq 3} (15-2x) \cdot P(X=x) \\ &= 15 \cdot \frac{2^0 e^{-2}}{0!} + 13 \cdot \frac{2^1 e^{-2}}{1!} + 11 \cdot \frac{2^2 e^{-2}}{2!} + 9 \cdot \frac{2^3 e^{-2}}{3!} = 75e^{-2} \end{aligned}$$

Note: A common mistake students make is to apply linearity of expectation in this case. In other words, they compute expectation of R using

$$\mathbb{E}[R] = \mathbb{E}[15-2X] = 15 - 2\mathbb{E}[X]$$

However, the issue with this approach is that R does not always equal $15-2X$. When $X > 3$, R actually equals to 0, making the first equality does not hold for all possible value of X. In other words, $\mathbb{E}[X]$ takes expectation over X for all values, including those with $X > 3$, but for those X, R should be 0 instead of $15-2X$.

In general, for this problem, there is no good way to compute the correct result using linearity of the expectation due to the fact that $R(x)$ is a piece-wise function.

c. (6 points) The baby wakes up half way through the night. What is the probability that she will not wake up again?

As the baby has woken up half way through the night, the remaining time available for the baby to wake up is the second-half of the night, which takes 4 hours as described by the question.

Hence, the question can be translated into asking what is the probability that the baby wakes up in the next 4 hours. This can be modeled as a Poisson distribution again. Notice that waking up during the first half of the night does not influence the rate of event waking up occurring. Therefore, given this condition, let X be the number of times the baby wakes up in **next 8 hours** is still following Poi(2). However, notice that know we are interested in Y, representing the number of times the baby wakes up in the **next 4 hours**. Hence, $Y \sim$ Poi(1) instead.

Because we are trying to compute that the baby never wake up in the second-half of the night, we are interested in $P(Y=0) = \frac{\lambda^0 e^\lambda}{0!} = e^{-1}$

d. (12 points) Let X_1 be the first time she wakes up, measured in hours after she went to sleep. Given that she wakes up exactly once in the night, $N = 1$, prove mathematically that $X_1 \sim$ Uni(0, 8).
Hint: either show that $P(X_1 < x|N = 1) = \frac{x}{8}$ or that $f(X_1 = x|N = 1) = \frac{1}{8}$. Both are equally hard.

Method 1:
Proof for $f(X_1 = x|N = 1) = \frac{1}{8}$.
Start with:

$$f(X_1 = x|N = 1) = \frac{P(N = 1, X_1 = x)}{P(N = 1)}$$

To find $P(N = 1, X_1 = x)$, note that there are two independent events here: the baby waking up exactly once at X_1, and the baby not waking up in the remaining 8-x hours.
To find the probability of the former event is to find the probability that X_1 exactly equals x.
We note that the distribution of X is the distribution of hours until the first time the baby wakes up. The description aligns well with the Exponential distribution, with parameter $\lambda = \frac{1}{4}$. λ is calculated because the rate is 2 per 8 hours, so $\frac{1}{4}$ per hour.
So,

$$f(X_1 = x) = \frac{1}{4} \cdot e^{-\frac{x}{4}}$$

Then, we find the probability that the baby doesn't wake up in the remaining 8-x hours. The Poisson parameter is $\lambda = 2$ for an 8-hour interval, so for an (8-x)-hour interval, the parameter would be $\lambda = 2 \cdot \frac{8-x}{8} = 2 - \frac{x}{4}$.
Let Y denote the number of times the baby wakes up in the 8-x hours. $Y \sim \text{Poisson}(2 - \frac{x}{4})$.
Hence,

$$P(Y = 0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-2+\frac{x}{4}}$$

Multiply the above two to find $P(N = 1, X_1 = x)$:

$$P(N = 1, X_1 = x) = P(Y = 0) \cdot f(X_1 = x)$$

Lastly, $P(N = 1)$ comes from 4a and is $2e^{-2}$.
Wrapping it up,

$$f(X_1 = x|N = 1) = \frac{P(Y = 0) \cdot f(X_1 = x)}{P(N = 1)} = \frac{e^{-2+\frac{x}{4}} \cdot (\frac{1}{4} \cdot e^{-\frac{x}{4}})}{2e^{-2}} = \frac{1}{8}$$

Method 2:
Proof for $f(X_1 < x|N = 1) = \frac{x}{8}$.
Start with:

$$P(X_1 < x|N = 1) = \frac{P(N = 1, X_1 < x)}{P(N = 1)}$$

Finding $P(N = 1, X_1 < x)$ is tricky. It requires that the baby wakes up once over x hours, and not waking up after the x hours. So there are two independent events here: waking up during the first x hours, and not waking up during the remaining 8-x hours. We find their respective probabilities and multiply them.
Let Z denote the number of times the baby wakes up over x hours. $Z \sim Poisson(\lambda = \frac{x}{4})$, because the rate is 2 per 8 hours and $\frac{x}{4}$ per x hours.
Then, you may find the probability that the baby doesn't wake up in the remaining 8-x hours as in Method 1 and is $e^{-2+\frac{x}{4}}$.
Wrapping it up,

$$P(X_1 < x|N = 1) = \frac{P(N = 1, X_1 < x)}{P(N = 1)} = \frac{\frac{x}{4} \cdot e^{-\frac{x}{4}} \cdot e^{-2+\frac{x}{4}}}{2e^{-2}} = \frac{x}{8}$$

5 Longest Sequence of Heads [20 points]

We are going to define a new random variable type $X \sim \text{Run}(n)$ where $P(X = x)$ is the probability of getting a **longest run** of exactly x **heads** in n coin flips. Here are a few examples with $n = 8$ coin flips:

$HHHHHHHH$	$X = 8$ since they are all heads.
$HTHTHTHT$	$X = 1$ since there are no consecutive runs with two heads.
$HTTHHHHT$	$X = 4$ since the longest run has 4 heads (HHHH).
$HHTTTTHH$	$X = 2$ since the longest run has 2 heads (HH).

a. (2 points) Let $X \sim \text{Run}(n)$. What is the support of X (the values that X can take on) in terms of n ?

0 through n

b. (6 points) Let $X \sim \text{Run}(n = 3)$. Provide the full Probability Mass Function for X:

Possible outcomes:

$X = 0$	TTT
$X = 1$	HTT, TTH, THT, HTH
$X = 2$	HHT, THH
$X = 3$	HHH

$P(X = 0,1,2,3) = \frac{1}{8}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$

Common mistakes:

- Some iteration of $\binom{3}{k} * (.5)^3$. This will include HTH in X=2 because out of 3 coins, you choose two to be heads. Because these heads are not consecutive, HTH cannot be included in getting a run of 2 heads. This PMF results in incorrect probabilities for X = 1 and X = 2.
- $P(X = 0,1,2,3) = \frac{1}{8}, \frac{3}{8}, \frac{2}{8}, \frac{1}{8}$. Note that this PMF does not sum to 1, so this cannot be your answer.

c. (6 points) You are given a function $\texttt{count2}(n)$ which returns the count of all unique orderings of n coin flips that have a longest run of 2 or fewer heads.

Here are a few base case examples:
$\texttt{count2}(0) = 2^0 = 1$, by definition.
$\texttt{count2}(1) = 2^1 = 2 \quad \{H, T\}$
$\texttt{count2}(2) = 2^2 = 4 \quad \{HH, HT, TH, TT\}$

Let $X \sim \text{Run}(n = 8)$. Define the probability $P(X \leq 2)$ using the `count2` function.

$\texttt{count2}(8)$ gives us exactly the event space that we want. The sample space is all sequences of 8 coin flips, of which there are 2^8. So $P(X \leq 2) = \frac{\texttt{count2}(8)}{2^8}$.

d. (6 points) What is $\texttt{count2}(n)$ if $n > 2$? You may define your formula using a recursive call to count where the parameter is a value less than n, for example $\texttt{count2}(n-2)$.
Hint: consider sequences that start with T, HT, and HHT.

How many length n sequences are there that begin with T and have a longest run of 2 or fewer heads? We know the first position has to be T, but then we have the remaining substring of $n-1$ positions to work with. Because the character that comes before this substring is a T, we don't need to worry about it influencing the maximum run length of heads within our substring. So the number of valid length $n-1$ substrings – and therefore the number of valid length n strings beginning with T – is exactly `count2(n-1)`.

Therefore the counts of valid sequences that start with T, HT, and HHT, respectively, are $\texttt{count2}(n-1)$, $\texttt{count2}(n-2)$, and $\texttt{count2}(n-3)$. Since these are all mutually exclusive, we have $\boxed{\texttt{count2}(n) = \texttt{count2}(n-1) + \texttt{count2}(n-2) + \texttt{count2}(n-3)}$.

We checked any other formulas that you provided, but unfortunately it's easy to find something that works for small cases and fails for larger ones. The only equivalent form we saw was $\texttt{count2}(n) = 2\texttt{count2}(n-2) + 2\texttt{count2}(n-3) + \texttt{count2}(n-4)$, which is what you get if you take the original answer and expand out the $\texttt{count2}(n-1)$ part according to the same formula. This has the small issue that $\texttt{count2}(n-4)$ may be undefined (since the problem is only guaranteeing that $n > 2$).

That's all folks! The goal of the Doodle Polls, and Color.com problems was to inspire new ways to see probability in your day to day life. Non-invasive ways to measure babies ability to hear is an open probability problem (and one that we might make some progress on). Getting rest as a new parent is real. Problem (5), Longest sequence of heads, is a step towards solving the challenge: given a poisson process of baby wake ups, what the the probability of getting a chunk of sleep at least k units long.